

# Oberon

**Voorspellende waarde van LOVS toetsen voor  
schoolsucces**



**Kennisnet**

het onderzoek is uitgevoerd met financiering van  
Kennisnet

# **Voorspellende waarde van LOVS toetsen voor schoolsucces**

Opdrachtgever: Stichting Kennisnet

Utrecht, februari 2013

Drs. E. van Aarsen (Oberon) m.m.v. Dr. J. Roeleveld (Kohnstamm Instituut / Universiteit van Amsterdam) en Dr. J.W. Luyten (Universiteit Twente)

© Oberon

Postbus 1423

3500 BK Utrecht

tel. 030-2306090

fax 030-2306080

e-mailadres: [info@oberon.eu](mailto:info@oberon.eu)

## Inhoudsopgave

<b>1</b>	<b>Aanleiding voor dit onderzoek .....</b>	<b>3</b>
<b>2</b>	<b>Onderzoeksaanpak.....</b>	<b>5</b>
<b>3</b>	<b>Relatie tussen toetsscores en het schooladvies.....</b>	<b>6</b>
<b>4</b>	<b>Voorspellende waarde van LOVS-toetsen voor de positie in het vierde jaar van het VO ....</b>	<b>10</b>
<b>5</b>	<b>Conclusies.....</b>	<b>13</b>
	<b>Literatuur .....</b>	<b>14</b>
<b>Bijlage 1</b>	<b>Methodologische keuzes .....</b>	<b>15</b>
<b>Bijlage 2</b>	<b>Voorspellende waarde door de jaren heen .....</b>	<b>17</b>
<b>Bijlage 3</b>	<b>Ongestandaardiseerde betacoëfficiënten.....</b>	<b>18</b>

# 1 Aanleiding voor dit onderzoek

## Totstandkoming van het schooladvies

Het Nederlandse basisonderwijs kent geen centrale examinering, zoals het voortgezet onderwijs dat bijvoorbeeld wel kent. Wel wordt aan het einde van de basisschool voor elke leerling een schooladvies opgesteld dat voor een groot deel bepaalt voor welk vervolgonderwijs de leerling in aanmerking komt. Voor het opstellen van dit advies zijn geen wettelijke regels vastgelegd. Scholen zijn in principe vrij om zelf te bepalen hoe zij het schooladvies opstellen. Onlangs is met het Wetsvoorstel voor verplichte invoering van de Eindtoets PO<sup>1</sup> gepleit voor het invoeren van een voor iedere leerling verplichte eindtoets, maar dit voorstel heeft expliciet niet tot doel de totstandkoming van het schooladvies verder in te kaderen. Paradoxaal genoeg wordt met het voorstel beoogd de invloed van het schooladvies meer centraal te stellen voor de toelating op het VO ten opzichte van de Eindtoetsscore.<sup>2</sup>

De afgelopen jaren zijn diverse onderzoeken uitgevoerd naar de totstandkoming van schooladviezen. Zo blijkt uit vragenlijstonderzoek onder directies en leerkrachten van basisscholen, dat cognitieve competenties<sup>3</sup> het sterkst meewegen bij het opstellen van het advies (Driessen en Doesborgh, 2005), maar dat ook sociaal-emotionele capaciteiten<sup>4</sup> behoorlijk sterk meewegen in het advies. Deze capaciteiten worden door de directie en leerkrachten bijna net zo belangrijk gevonden als de cognitieve capaciteiten (Driessen, 2005). De ondersteuning die kinderen van huis uit krijgen voor hun schoolwerk en hun beheersing van de Nederlandse taal blijken minder bepalend voor het advies. Onderzoek van Driessen (2011) laat verder zien dat, rekening houdend met competenties en achtergronden, er binnen grote gemeenten niet anders wordt geadviseerd dan in kleine. Uit andere studies blijkt echter dat niet-cognitieve competenties<sup>5</sup> maar weinig bijdragen aan het advies; in hun studie constateren Luyten en Bosker (2004) dat ongeveer 70% van de variantie in schooladviezen verklaard wordt door toetsprestaties, en dat niet-cognitieve competenties hierop nauwelijks aanvullende verklaringskracht bieden.

## Behoeftte aan een eenvoudige indicator van schoolvorderingen naast of in plaats van de Eindtoets

Hoewel VO-scholen in merendeel het advies van het PO volgen<sup>6</sup>, merken we in gesprekken met betrokkenen bij VO-scholen behoefte aan een meer transparante onderbouwing van het advies. In de overdracht van leerlingen worden soms wel toetsscores uit het leerlingvolgsysteem (LOVS-scores) van het basisonderwijs meegestuurd, maar deze blijken voor het VO vaak moeilijk te interpreteren. Vaak is onduidelijk wat deze scores betekenen, hoe ze zijn meegewogen in het advies, en hoe groot hun zeggingskracht is voor later schoolsucces. Doordat het schooladvies voor VO-scholen vaak een 'black box' is, hanteren veel VO-scholen momenteel toelatingscriteria op basis van de Cito-Eindtoetsscore. Deze praktijk ligt echter onder vuur<sup>7</sup> en is voor veel scholen onhoudbaar in het scenario waarin de Eindtoets later in het schooljaar wordt afgenomen. Onder meer in Amsterdam, Almere en de regio rondom de stad Utrecht werkt men inmiddels aan of al met een methodiek waarin de schoolvorderingen zoals die kunnen worden afgelezen uit toetsscores uit het leerlingvolgsysteem een grote rol krijgen bij de toelating tot VO-scholen.

<sup>1</sup> Zie: <http://www.rijksoverheid.nl/bestanden/documenten-en-publicaties/kamerstukken/2012/01/31/voorstel-van-wet-in-verband-met-de-invoering-van-een-centrale-eindtoets/voorstel-van-wet.pdf>

<sup>2</sup> Door deze eindtoets later in het jaar te laten afnemen, is de veronderstelling dat de meeste leerlingen zich al hebben aangemeld en zijn toegelaten tot een VO-school, waardoor de Eindtoets een minder prominente rol zou krijgen in het toelatingsbeleid van scholen.

<sup>3</sup> toetsprestaties, capaciteiten en talent

<sup>4</sup> interesses, doorzettingsvermogen en de sociaal-emotionele stabiliteit van een leerling

<sup>5</sup> Zoals de werkhouding van leerlingen, hun autonomie, ordelijkheid en de betrokkenheid van ouders

<sup>6</sup> In ons databestand is 91% van de leerlingen in het VO gestart in een onderwijstype dat in lijn ligt met hun schooladvies. Hierbij speelt mee dat veel leerlingen starten in combinatieklassen en een (klein) deel (7%) van de leerlingen in ons bestand een combinatie-advies heeft gekregen.

<sup>7</sup> zie o.m. het Wetsvoorstel voor verplichte invoering van de Eindtoets PO, p24

In opdracht van Kennisnet heeft Oberon met medewerking van Dr. J. Roeleveld (Kohnstamm Instituut / Universiteit van Amsterdam) en Dr. J.W. Luyten (Universiteit Twente) een onderzoek uitgevoerd naar de relatie tussen LOVS-scores en schooladvies en hun voorspellende waarde op later schoolsucces in het VO.

## 2 Onderzoeksaanpak

In dit onderzoek staan de volgende vragen en subvragen (schuingedrukt) centraal:

1. Hoe sterk is de **relatie tussen LOVS-toetsen en het schooladvies**; *Hoe wegen de domeinen begrijpend lezen, rekenen-wiskunde, en spelling hierin mee, en welk gewicht krijgen de toetsen uit groep 8, 7 en 6? Hoe zit dit voor de Cito-Eindtoets?*
2. Wat is de **voorspellende waarde van LOVS-toetsen voor de positie in het vierde jaar van het VO**? *Hoe wegen de domeinen begrijpend lezen, rekenen-wiskunde, en spelling hierin mee, en welk gewicht krijgen de toetsen uit groep 8, 7 en 6? En welk gewicht krijgt het advies zelf? En hoe zit dit voor de Cito-Eindtoets? Met welke berekeningswijze bieden toetsscores de beste voorspelling van positie in het vierde jaar van het VO?*

Voor dit onderzoek hebben we statistische analyses verricht op toetsgegevens uit de groepen 6, 7 en 8 van het basisonderwijs en de doorstroom van leerlingen tot en met hun vierde jaar in het VO. De gebruikte gegevens zijn afkomstig van cohorten die gestart zijn in het VO tussen 2006 en 2012. In bijlage 1 geven we een toelichting op de methodologische keuzes die we in dit onderzoek hebben gemaakt.

In totaal hebben we voldoende scores van 10.438 leerlingen, verspreid over 124 basisscholen / locaties.<sup>8</sup> Van deze leerlingen is in ruim 70% van de gevallen bekend welk schooladvies zij hebben gekregen. 23% van de leerlingen kan met de huidige data worden gevolgd tot en met hun vierde jaar in het VO; de leerlingen zitten dan verspreid over 19 VO-locaties. Hoewel de gegevens afkomstig zijn van scholen in grote gemeenten, levert dit geen beperking op van de validiteit van de uitkomsten (vergelijk Driessen, 2011).

### Toetsversies

De afgelopen jaren zijn nieuwe LOVS-toetsen op de markt gekomen die inmiddels op vrijwel alle scholen in dit onderzoek worden afgenomen. Omdat deze toetsen niet eenvoudig onderling vergelijkbaar zijn, hebben we de analyses voor beide toetssystemen afzonderlijk gedaan. Voor de analyses over de doorstroom in het VO kunnen überhaupt alleen de oude toetsen worden meegewogen, omdat de leerlingen die de nieuwe toetsen hebben afgenomen pas vanaf 2011/12 zijn ingestroomd in het VO en dus nog geen vier jaar in het VO zitten.

---

<sup>8</sup> Waarvan sommige scholen met maar weinig leerlingen met genoeg toetsscores.

### 3 Relatie tussen toetsscores en het schooladvies

Scores op de LOVS-toetsen afgenomen in groep 6 tot en met 8 geven een heel hoge voorspelling van het advies dat de leerling krijgt. De proportie verklaarde variatie ( $R^2$ ) bedraagt als alle toetsen worden meegewogen .761 voor de oude LOVS-toetsen en .817 voor de nieuwe toetsen. Overigens bieden “zuiniger” modellen, waarin alleen een deel van de toetsen wordt meegewogen ook al een goede voorspelling van het advies. In onderstaande tabel hebben we verschillende varianten van het model weergegeven, met hun proportie verklaarde variantie. Bij de nieuwe LOVS-toetsen bieden de toetsscores een nog betere voorspeller van het schooladvies van de leerling. Dit kan er op duiden dat de nieuwe toetsen meer onderscheidend zijn, of dat leerkrachten bij het opstellen van het advies tegenwoordig meer afgaan op toetsscores en minder op overige zaken.

Tabel 3.1. Proportie verklaarde variantie PO-toetsen voor schooladvies (oude en nieuwe toetsen)

Model	$R^2$	$R^2$
	Oude toetsen	Nieuwe toetsen
1. BL en RW groep 8	.706	.760
2. Bovenstaande + SP groep 8	.725*	.783*
3. Bovenstaande + BL, RW en SP groep 7	.753*	.813*
4. Bovenstaande + BL, RW en SP groep 6	.761*	.817*
	N= 5532	N=1812
5. Alleen Cito-Eindtoets	.705	.764 <sup>9</sup>
	N= 3947	N=1466

\* de toegevoegde toetsen zorgen voor een significante toename van de proportie verklaarde variantie t.o.v. het voorgaande model

Een deel van de leerlingen heeft ook de Cito-Eindtoets afgenomen.<sup>10</sup> Bij deze leerlingen is te zien dat de score op de Eindtoets ook een goede voorspelling geeft van het schooladvies (model 5), maar net iets minder goed dan de scores op de verschillende LOVS-toetsen gezamenlijk. Overigens zien we bij leerlingen die de Eindtoets afnemen dat hun LOVS-scores het advies minder goed verklaren dan bij leerlingen die de Eindtoets niet afnemen. Hierbij kan meespelen dat scholen de Eindtoets laten meewegen bij het advies, en daardoor de LOVS-toetsen meer op de achtergrond plaatsen.

#### Vertaling naar de praktijk

Als we met de berekeningwijze die uit bovenstaand model (4) vloeit, enkel op basis van de LOVS-toetsscores een advies moesten formuleren (zonder de leerling en diens ambities, werkhouding, etc. te kennen) zouden we in 60% van de gevallen het zelfde advies geven als de leerkracht.<sup>11</sup> In 19% van de gevallen zouden we één niveau hoger adviseren, en in 20% van de gevallen één niveau lager (bij 2% van de leerlingen zouden we meer dan 1 niveau afwijken).<sup>12</sup>

Ons model wijkt met name af van het leerkrachtadvies bij leerlingen die door de leerkracht het advies BBL hebben gekregen. Relatief veel BBL-leerlingen zouden door ons, op basis van hun scores het advies VMBO-KL krijgen. In onderstaande tabel is dit te zien.

<sup>9</sup> Het gaat hier om de Cito-Eindtoetsscore van leerlingen die ook de nieuwe LOVS-toetsen hebben afgenomen. Immers de opgaven en normering van de Cito-Eindtoets worden continu vernieuwd.

<sup>10</sup> Van circa 74% van de leerlingen met een bekend advies is ook een Eindtoetsscore bekend

<sup>11</sup> De leerlingen die een combinatieadvies hebben gekregen, zijn hierbij buiten beschouwing gelaten. Als we die wel meetellen, krijgt 62% van de leerlingen het zelfde advies als berekend.

<sup>12</sup> De Cohen's Kappa bedraagt .488 voor leerlingen met de oude LOVS-toetsen. De nieuwe toetsen schatten 65% van de leerlingen hetzelfde als de leerkracht. De Cohen's Kappa bedraagt .554

Tabel 3.2. Overeenkomst berekend advies o.b.v. LOVS en schooladvies (oude toetsen)

Leerkracht	VMBO -BL	VMBO -KL	VMBO -GL/TL	HAVO	VWO
<b>Berekend obv LOVS (oude toets)</b>					
VMBO -BL	<b>37%</b>	5%	1%	0%	0%
VMBO -KL	60%	<b>70%</b>	26%	2%	0%
VMBO -GL/TL	3%	24%	<b>60%</b>	28%	1%
HAVO	0%	1%	13%	<b>60%</b>	28%
VWO	0%	0%	1%	10%	<b>70%</b>
Totaal	100%	100%	100%	100%	100%
	766	821	1398	1323	850

Ter vergelijking, als we het advies zouden schatten op basis van de Cito-Eindtoets (model 5) zouden we bij 63% van de leerlingen op het zelfde advies uitkomen als de leerkracht.<sup>13</sup> De overeenkomst tussen de Cito-uitkomst en het advies is met name groot (groter dan het LOVS-model) bij leerlingen met leerkrachtadvies Havo, en in mindere mate ook bij leerlingen met advies VMBO-GL/TL. Daarentegen is de overeenkomst tussen Eindtoetsscore en advies weer een stuk kleiner dan bij het LOVS-model als het gaat om de adviezen VMBO-BL en -KL en VWO.<sup>14</sup>

Tenslotte valt op dat de deelname aan de Eindtoets erg laag is onder leerlingen met het advies VMBO-BL. Minder dan 7% van hen heeft een Eindtoetsscore, terwijl dit voor de overige adviezen loopt van 72% (KL) tot 90% (VWO).<sup>15</sup>

Tabel 3.3. Overeenkomst berekend advies o.b.v. Cito-Eindtoets en schooladvies (leerlingen met oude toetsen)

Leerkracht	VMBO -BL	VMBO -KL	VMBO -GL/TL	HAVO	VWO
<b>Berekend obv Eindtoets (leerlingen met oude LOVS-toets)</b>					
VMBO -BL	<b>27%</b>	3%	0%	0%	0%
VMBO -KL	67%	<b>51%</b>	12%	1%	0%
VMBO -GL/TL	6%	46%	<b>63%</b>	19%	0%
HAVO	0%	0%	24%	<b>73%</b>	41%
VWO	0%	0%	0%	8%	<b>58%</b>
Totaal	100%	100%	100%	100%	100%
	51	595	1168	1183	762

### Weging van de verschillende domeinen en meetmomenten van de LOVS-toetsen

Uit onze analyses blijkt dat de scores op rekenen-wiskunde en in mindere mate begrijpend lezen in groep 8 het sterkst samenhangen met het advies. Scores op spelling en toetsen uit eerdere leerjaren wegen maar beperkt mee in het advies, maar bieden in combinatie met de andere toetsen toch een (iets) betere benadering van het advies dan zonder deze toetsen (zie ook het model op de vorige pagina).

<sup>13</sup> Het gaat hierbij om 'adviezen' die op basis van de data zijn berekend. Niet te verwarren met de toetscriteria die door veel scholen worden gehanteerd in hun aanmeldingssysteem.

<sup>14</sup> Cohen's Kappa bedraagt .489

<sup>15</sup> Dit lijkt sterk samen te hangen met de gekozen cohorten: in 2012 is van 71% van de leerlingen met advies VMBO-BL de citoscore bekend.



Tussen de oude en nieuwe toetsen blijken op enkele punten aanzienlijke verschillen te bestaan in het relatieve gewicht dat een toets heeft binnen het berekende advies. Zo heeft de nieuwe toets begrijpend lezen in groep 7 een veel grotere samenhang met het advies dan de oude toets voor groep 7. In bijlage 3 geven we ook de ongestandaardiseerde betacoëfficiënten weer.

Tabel 3.4. Relatief gewicht van losse toetsen op berekend advies (oude en nieuwe toetsen)

Toets	Gestandaardiseerde Betacoëfficiënt Oude toetsen	Verhouding	Gestandaardiseerde Betacoëfficiënt Nieuwe toetsen	Verhouding
Rekenen-Wiskunde gr. 8	.337	31%	.349	32%
Begrijpend Lezen gr. 8	.176	16%	.195	18%
Rekenen-Wiskunde gr. 7	.126	12%	.076	7%
Begrijpend Lezen gr. 6	.116	11%	.065	6%
Begrijpend Lezen gr. 7	.111	10%	.189	17%
Spelling gr. 7	.073	7%	.069	6%
Spelling gr. 8	.065	6%	.062	6%
Rekenen-Wiskunde gr. 6	.051	5%	.059	5%
Spelling gr. 6	.041	4%	.043	4%

### De proef op de som: Werkt het ook voor subgroepen?

Het model en de berekeningswijze zijn gebaseerd op de hele onderzoeksgroep: jongens en meisjes, verschillende gemeenten en verschillende cohorten. Een goede test voor ons model is te kijken of het ook voor verschillende onderliggende groepen “werkt” d.w.z. een goede voorspeller geeft van het schooladvies. Als dat zo is, betekent dit dat het model naar verwachting vergelijkbare resultaten zal bieden bij leerlingen die niet in de onderzoeksgroep vallen.

In onderstaande tabellen geven we de proportie verklaarde variantie weer die het model biedt voor de verschillende subgroepen en de mate van samenhang tussen het berekende advies en het door de leerkracht gegeven advies. Hoewel er verschillen te zien zijn, bieden de toetsscores voor elk van de subgroepen een goede verklaring van het advies.

Tabel 3.5. Overeenstemming berekend advies o.b.v. LOVS met schooladvies voor subgroepen (oude en nieuwe toetsen)

		% advies <sup>16</sup> gelijk als berekend Oude toetsen	Cohen's Kappa Oude toetsen	% advies gelijk als berekend Nieuwe toetsen	Cohen's Kappa Nieuwe toetsen
Geslacht	Jongens	59%	.482	64%	.541
	Meisjes	60%	.494	66%	.567
Cohort	2006	62%	.498	-	
	2007	55%	.430	-	
	2008	58%	.473	-	
	2009	64%	.540	-	
	2010	60%	.487	-	
	2011	60%	.477	66%	.569
	2012	63%	.519	65%	.548

<sup>16</sup> In deze tabel zijn de gecombineerde adviezen buiten beschouwing gelaten. Als die wel meetellen, stijgt het percentage gelijke adviezen.

	<b>% advies<sup>17</sup> gelijk als berekend Oude toetsen</b>	<b>Cohen's Kappa Oude toetsen</b>	<b>% advies gelijk als berekend Nieuwe toetsen</b>	<b>Cohen's Kappa Nieuwe toetsen</b>
Gemeente A	60%	.487	65%	.556
B	61%	.492	65%	.545

<sup>17</sup> In deze tabel zijn de gecombineerde adviezen buiten beschouwing gelaten. Als die wel meetellen, stijgt het percentage gelijke adviezen.

## 4 Voorspellende waarde van LOVS-toetsen voor de positie in het vierde jaar van het VO

In het voorgaande hebben we een model beschreven waarin we proberen het advies van de school zo *dicht mogelijk te benaderen* op basis van scores op LOVS-toetsen. Op deze manier gaan we er impliciet van uit dat het advies van de basisschool het juiste is.

De vraag is of we op grond van de LOVS-scores (terugkijkend) een “*beter*” advies kunnen formuleren dan de leerkracht. In dit geval vatten we “*beter*” op als: een betere voorspelling van de positie in het vierde jaar van het VO. (Uiteraard kun je hierbij de discussie voeren of het advies slechts een indicatie moet geven van het brugklasniveau dat de leerling zou passen, of ook een verdergaande prognose moet bieden: op welk niveau kan de leerling op niveau doorstromen en uiteindelijk (onvertraagd) het diploma halen. Wij kiezen hierbij dus voor het laatste, verdergaande prognose).

### Voorspellende waarde van het advies alleen, en in combinatie met LOVS-toetsen

Van de bijna 2400 leerlingen die we ten opzichte van hun advies kunnen volgen tot en met hun vierde jaar in het VO ligt zo'n 55% op koers van hun schooladvies. 29% van de leerlingen is blijven zitten, of zit inmiddels op een lager niveau dan hun advies. 16% van de leerlingen zit op een hoger niveau. Het advies verklaart dan ook een groot deel van de positie in het vierde jaar ( $R^2 = .600$ ). Hierbij speelt wel mee dat het advies werkt als een “self-fulfilling prophecy”; het advies wordt meestal (91%) overgenomen door de VO-school, waardoor de leerling in de brugklas al voorgesorteerd is voor dat onderwijstype. De leerling hoeft dus geen 'barrières' te doorbreken (kiezen om een leerling een niveau hoger of lager te laten doorstromen, misschien zelfs naar een andere school te laten gaan) om in de vierde van dat onderwijstype te komen.

Tabel 4.1. *Proportie verklaarde variantie advies en LOVS voor positie op de leerjarenladder in het vierde VO-jaar (oude toetsen)*

Model	$R^2$
Oude toetsen	
1. Advies	.600
2. Bovenstaande + BL en RW groep 8	.626*
3. Bovenstaande + SP groep 8	.627
4. Bovenstaande + BL, RW en SP groep 7	.634*
5. Bovenstaande + BL en SP groep 6 <sup>18</sup>	.635*
N= 2393	

\* de toegevoegde toetsen zorgen voor een significante toename van de proportie verklaarde variantie t.o.v. het voorgaande model

Als we de LOVS-toetsen toevoegen, bovenop het advies, neemt de proportie verklaarde variantie van het model nog iets toe (wordt  $R^2 = .636$ ). Gegeven een bepaald advies dragen toetsscores dus bij aan een iets betere voorspelling van het niveau na 4 jaar. VO-scholen doen er al met al dus goed aan om zich te verdiepen in de scores die leerlingen op het PO hebben behaald, als zij risico-leerlingen, of talenten willen signaleren.

<sup>18</sup> De toets voor rekenen-wiskunde in groep 6 correleert in dit model (heel licht) negatief met de positie in het 4e jaar in het VO. Aangezien deze twee variabelen onderling (zonder alle overige variabelen) wel positief correleren, duidt dit op multicollineariteit; de score op Rekenen-Wiskunde gr. 6 hangt sterk samen met de overige variabelen, waardoor de relatie met de positie in het vierde jaar arbitrair wordt. Om deze reden hebben we de Rekenen-Wiskunde gr. 6 buiten ons model gelaten.

### Voorspellende waarde van LOVS-toetsen, en van de Cito-Eindtoets

Het doel van dit onderzoek is echter om onafhankelijk van het oordeel van de leerkracht, alleen op basis van LOVS-scores te komen tot een advies. Als we alleen de LOVS-scores bekijken, en het leerkrachtadvies negeren, kunnen we ook al een redelijk goede voorspelling geven van de positie in het vierde jaar van het VO. De maximale proportie verklaarde variatie ( $R^2$ ) bedraagt .575 (zie onderstaande tabel). De beste voorspelling wordt bereikt als we naast toetsen uit groep 8 ook die uit groep 7 en 6 meewegen. In de praktijk betekent dit dat onze voorspelling op basis van model 4 voor 44% van de leerlingen uitkomt.<sup>19</sup> In 21% van de gevallen schatten we de leerling één niveau of leerjaar te hoog, en in 25% van de gevallen schatten we één niveau of leerjaar te laag.

Tabel 4.2 Proportie verklaarde variantie PO-toetsen voor positie op de leerjarenladder in het vierde VO-jaar (oude toetsen)<sup>20</sup>

Model	$R^2$ Oude toetsen
1. BL en RW groep 8	.527
2. Bovenstaande + SP groep 8	.537*
3. Bovenstaande + BL, RW en SP groep 7	.568*
4. Bovenstaande + BL en SP groep 6	.575*
	N= 2417
5. Alleen Cito-Eindtoets	.497
6. Cito-Eindtoets + LOVS toetsen (uit model 4)	.521*
	N= 1582

\* de toegevoegde toetsen zorgen voor een significante toename van de proportie verklaarde variantie t.o.v. het voorgaande model

Ter vergelijking hebben we ook de voorspellingskracht van de Cito-Eindtoets bekeken (model 5 hierboven). De proportie verklaarde variantie ( $R^2$ ) bedraagt .497 op de positie in het vierde jaar van het VO.<sup>21</sup> Voor de leerlingen die zowel LOVS-toetsen als de Eindtoets hebben afgenomen, bieden deze toetsen tezamen een betere voorspeller van later schoolsucces dan alleen de Eindtoets of alleen de genoemde LOVS-toetsen (Model 6:  $R^2=.521$ ).<sup>22</sup>

Al met al kunnen we dus vaststellen dat het schooladvies de beste voorspelling biedt van de positie die een leerling in het 4<sup>e</sup> jaar van het VO heeft. De LOVS-toetsen bieden ook een redelijk goede voorspelling: minder goed dan het schooladvies, maar beter dan de Cito-Eindtoetsscore.

Bovenstaande modellen gaan er van uit dat de leerkracht in groep 8 terugkijkt en een zo passend mogelijk advies formuleert. Andersom is het ook interessant om te bekijken hoe goed een leerkracht in groep 6 al zou kunnen voorspellen, en wat de scores in de daaropvolgende jaren toevoegen. In bijlage 2 gaan we daar op in.

### Weging van de verschillende domeinen en meetmomenten van de LOVS-toetsen

<sup>19</sup> Enkele leerlingen zijn gedoubleerd en zitten nog in een combinatieklas (3 havo/vwo). Deze leerlingen zijn bij het berekenen van de Cohen's Kappa en het percentage overeenstemming buiten beschouwing gelaten. De Cohen's Kappa kan immers alleen worden berekend wanneer de categorieën beide variabelen (voorspelde positie en feitelijke positie) aan elkaar gelijk zijn.

<sup>20</sup> Elke groep van toegevoegde toetsen in het model (1 t/m 4) zorgt voor een statistisch significante toevoeging aan de proportie verklaarde variantie.

<sup>21</sup> Op basis van alleen de Cito-Eindtoetsscore schatten we voor 45% van de positie in het vierde jaar correct. In 20% van de gevallen schatten we met de Eindtoetsscore de leerling één niveau of leerjaar te hoog, en in 26% van de gevallen schatten we één niveau of leerjaar te laag.

<sup>22</sup> De proportie verklaarde variantie op basis van het LOVS en Eindtoets tezamen is lager dan die op basis van alleen het LOVS. Dit komt doordat in deze berekening alleen scores worden meegeteld van leerlingen die aan beide toetsen hebben deelgenomen. Doordat leerlingen met een BBL-advies in het verleden vaak niet deelnamen aan de Cito-Eindtoets, vallen ook hun (lage) LOVS-scores buiten deze berekeningswijze, hetgeen zorgt voor een kleinere variantie in LOVS-scores.

Net als voor het berekenen van een advies geldt voor het voorspellen van de schoolloopbaan in het VO dat de scores op rekenen-wiskunde en in mindere mate begrijpend lezen in groep 8 de grootste voorspellingskracht bieden. Wel zien we dat de scores op begrijpend lezen zwaarder meewegen in de voorspelling van de positie in het 4e VO-jaar dan in het advies (vergelijk paragraaf 3.1). In bijlage 3 geven we ook de ongestandaardiseerde betacoëfficiënten weer.

Tabel 4.3. Relatief gewicht van losse toetsen op positie op de leerjarenladder in het vierde VO-jaar (oude toetsen)

Toets	Gestandaardiseerde Betacoëfficiënt	Verhouding
Rekenen-Wiskunde gr. 8	.249	26%
Begrijpend Lezen gr. 8	.195	20%
Begrijpend Lezen gr. 7	.144	15%
Begrijpend Lezen gr. 6	.118	12%
Rekenen-Wiskunde gr. 7	.109	11%
Spelling gr. 7	.087	9%
Spelling gr. 6	.038	4%
Spelling gr. 8	.012	1%

### De proef op de som: Werkt het ook voor subgroepen?

In de verschillende subgroepen komt ons model op basis van LOVS-toetsen tot vergelijkbare resultaten (zie onderstaande tabel). De grootste verschillen zijn te vinden tussen de verschillende cohorten. Omdat het aantal cases daar het meest beperkt is (het aantal leerlingen is verspreid over vier cohorten) is deze uitsplitsing ook het meest gevoelig voor uitschieters. Een inhoudelijke verklaring voor deze afwijkingen hebben we evenwel niet.

Tabel 4.5. Overeenstemming berekende positie o.b.v. LOVS met positie op de leerjarenladder in het vierde VO-jaar voor subgroepen (oude toetsen)

		% positie gelijk als berekend	% positie 1 niveau hoger dan berekend	% positie 1 niveau lager dan berekend
Geslacht	Jongens	43%	22%	25%
	Meisjes	45%	28%	18%
Cohort	2006	48%	22%	23%
	2007	41%	29%	18%
	2008	46%	22%	23%
	2009	41%	26%	25%
Gemeente	A	43%	25%	21%
	B	45%	24%	22%

## 5 Conclusies

### Relatie tussen toetsscores en het schooladvies

Uit ons onderzoek blijkt dat het schooladvies sterk samenhangt met de scores op de toetsen uit het Leerling- en Onderwijsvolgsysteem (LOVS). De scores op toetsen voor rekenen-wiskunde, begrijpend lezen en spelling in groep 6 tot en met 8 verklaren samen zo'n 76% van de variantie in adviezen. Bij de nieuwe LOVS-toetsen is dit zelfs bijna 82%. Deze grote samenhang maakt het mogelijk om op basis van toetsscores een "advies" te berekenen. Ter vergelijking: de Cito-Eindtoetsscore verklaart zo'n 70% van de variantie in adviezen (76% bij leerlingen die de nieuwe LOVS-toetsen hebben afgenomen). Dat betekent dat we, zonder de leerling, diens ambities, werkhouding, etc. te kennen, op basis van diens LOVS-scores een advies schatten dat in zo'n 60% van de gevallen (bij de nieuwe toetsen zelfs in 65% van de gevallen) gelijk is aan het (enkelvoudige) advies van de leerkracht.

VO-scholen kunnen deze kennis gebruiken om in te schatten welke leerlingen, ten opzichte van hun toetsscores relatief hoog of relatief laag zijn geadviseerd. Zo kan, bijvoorbeeld in overleg met de basisschool, de VO-school extra inzetten op leerlingbegeleiding.

### Voorspellende waarde van LOVS-toetsen voor de positie in het vierde jaar van het VO

Uit dit onderzoek blijkt eveneens dat de schoolloopbaan in de eerste jaren van het VO een stuk moeilijker te voorspellen is dan het advies. Desondanks kan ruim 57% van de variantie in de positie op de leerjarenladder van de leerling in het vierde jaar van het VO worden verklaard op basis van LOVS-toetsen. De LOVS-scores bieden daarmee een iets minder goede voorspeller dan het leerkrachtadvies (zo'n 60%), maar wel weer beter dan de Cito-Eindtoets (zo'n 50%).

Het leerkrachtadvies biedt al met al dus de beste voorspeller van later schoolsucces, waarbij opgemerkt dient te worden dat dit advies ook vrijwel altijd (91%) wordt overgenomen door de VO-school, en dus als een '*self-fulfilling prophesy*' kan werken.

### Vervolg

De inzichten uit dit onderzoek en de in dit onderzoek toegepaste berekeningswijzen kunnen in de toekomst worden verwerkt in leerlingvolgsystemen; ofwel als handvat voor leerkrachten bij het formuleren van het advies, ofwel als onafhankelijk gegeven waartegen het VO het leerkrachtadvies kan afzetten. Bij deze implementatie kan worden gekozen voor een *uitgebreid model* of een meer *zuinig model*. Het uitgebreide model, waarbij ook scores uit groep 7 en 6 worden meegewogen, biedt een iets betere voorspelling van het advies en later schoolsucces en "dempt" beter voor uitschieters (hoge of lage scores). Het zuinige model kan ook worden toegepast bij leerlingen waarvan geen scores uit groep 6 en 7 bekend zijn, zonder dat deze moeten worden geschat op basis van de toetsen die de leerling wel heeft gemaakt.

Het feit dat de relatie tussen LOVS-scores en advies bij de nieuwe toetsen (en nieuwe cohorten) een stuk hoger is dan voor de oude toetsen en cohorten biedt perspectief voor de voor het voorspellen van schoolsucces van meer recente cohorten. Het valt immers te verwachten dat deze toetsen ook een betere indicator geven van later schoolsucces dan de oude toetsen.

## Literatuur

- Bosker, R.J., Velden, R.K.W. van der, & Hofman, W.H.A. (1985). *Een generatie geselecteerd. Deel 1: De loopbanen*. Groningen: RION.
- Driessen, G., & Doesborgh, J. (2005). *Relaties tussen achtergrondkenmerken en competenties van leerlingen en hun advies voor voortgezet onderwijs*. In G. Driessen, J. Doesborgh, G. Ledoux, M. Overmaat, J. Roeleveld, & I. van der Veen, *Van basis- naar voorgezet onderwijs. Voorbereiding, advisering en effecten* (pp. 39-70). Nijmegen/ Amsterdam: ITS/SCO-Kohnstamm Instituut.
- Driessen, G. (2011). *Onderadvisering van allochtone leerlingen?*. Nijmegen: ITS.
- Huisman, M. (2007). *Multipale imputatie van ontbrekende scores*. In A.E. Bronner, P. Dekker, E. de Leeuw, L.J. Paas, K. de Ruyter, A. Smidts, and J.E. Wieringa (Eds.), *Ontwikkelingen in het marktonderzoek, Jaarboek 2007 MarktOnderzoekAssociatie* (pp. 171-188). Haarlem: Spaar en Hout.
- Luyten, H., & Bosker, R. (2004). *Hoe meritocratisch zijn schooladviezen?* Pedagogische Studiën, 81, (2), 89-103.

## Bijlage 1 Methodologische keuzes

### Keuze voor toetsdomeinen

We hebben gekozen de medio-toetsen voor begrijpend lezen, rekenen-wiskunde en spelling mee te nemen in ons onderzoek. Graag hadden we de toetsen voor woordenschat hieraan toegevoegd, maar deze bleken onvoldoende vaak afgenomen.

### Selectie van toetsscores

Van elke leerling hebben we een toetsscore opgezocht die behaald is op de medio-toets op de domeinen begrijpend lezen, rekenen-wiskunde en spelling in groep 6, in groep 7 en groep 8. Alleen scores die halverwege het schooljaar behaald zijn (januari t/m maart) telden mee.<sup>23</sup> In gevallen waar de leerling de toets meerdere keren heeft afgenomen, telde alleen de hoogst behaalde score mee. Uiteindelijk zijn alleen de leerlingen meegerekend waarvan van alledrie de domeinen, twee of alledrie de scores per domein bekend zijn (bijvoorbeeld een score uit groep 6 en groep 8).

### Omzetten van toetsscores in gestandaardiseerde scores

De scores van de voor dit onderzoek gebruikte toetsen van Cito worden uitgedrukt in vaardigheidsscores. Voor elk van de domeinen geldt een eigen schaal, waardoor scores op het ene domein niet vergelijkbaar zijn met die op een ander domein. Wel zijn alle scores uit te drukken in een niveau-indicatie (A t/m E en voor de nieuwe toetsen ook een alternatieve indeling I t/m V). De niveaus geven aan hoe een leerling scoort in vergelijking met het landelijk niveau van leerlingen in dat leerjaar. Een belangrijk nadeel van die methode is dat hij erg grofmazig is: niveau B zegt bijvoorbeeld dat een leerling behoort tot de 50% beste leerlingen van Nederland, maar in werkelijkheid ligt de score 'ergens' tussen de 50% en 26% best scorende leerlingen. Hoewel Cito uit commerciële overwegingen geen inzage geeft in de exacte verdeling van toetsscores, kan op grond van de grenswaarden van elk niveau voor elke toetsscore een schatting worden gemaakt van de positie *binnen* elk niveau (ligt de score bijvoorbeeld dicht bij de bovenkant van B, dicht bij de onderkant of ergens tussenin).<sup>24</sup>

### Imputeren van ontbrekende toetsscores

Hoewel er voldoende cases (leerlingen) met een score op elke toets in alle drie de leerjaren in de dataset aanwezig waren, is er voor gekozen om voor leerlingen waar één score per domein mistte deze score te schatten. Wij hebben dit gedaan door middel van imputatie van de scores, op basis van lineaire regressie waarin de te schatten score wordt opgenomen als afhankelijke variabele, en de overige scores van dat domein als onafhankelijke variabelen. Als voorbeeld: De score op de rekentoets in groep 8 wordt dan geschat op basis van de scores die (qua scores) vergelijkbare leerlingen halen op de toetsen in groep 6 en 7. Aangezien de toetsscores in groep 6, 7 en 8 onderling hoog correleren, geeft dit een betrouwbare schatting. Hoewel deze methode enkele nadelen heeft ten opzichte van meer complexe multi-pele imputatiemethoden (zie bijvoorbeeld Huisman, 2007), is een groot voordeel dat de geschatte scores met eenvoudige formules te berekenen zijn, en daarmee geschikt zijn voor implementatie in leerlingvolgsystemen.

In onderstaande tabel is te zien hoeveel scores (en percentage van het totaal aantal leerlingen) bekend waren, en hoe groot het aantal geïmputeerde scores is.

<sup>23</sup> Met uitzondering van de leerlingen waarvoor een alternatieve toetsversie is ingezet betreft het de medio-toets behorend bij het leerjaar. Ook deze alternatieve toetsversies kunnen op dezelfde wijze worden genormeerd als de toets behorend bij het leerjaar.

<sup>24</sup> Het leerlingvolgsysteem ParnasSys maakt gebruik van een vergelijkbare methode, waarbij wij echter wel de nodige kanttekeningen plaatsen.



	Oude toetsen		Nieuwe toetsen	
	Score bekend voor imputatie	Scores geïmputeerd	Score bekend voor imputatie	Scores geïmputeerd
Begrijpend lezen gr 6	7246	650	2161	381
Begrijpend lezen gr 7	7580	316	2532	10
Begrijpend lezen gr 8	6677	1219	2040	502
Rekenen-wiskunde gr 6	6972	924	2295	247
Rekenen-wiskunde gr 7	7654	242	2532	10
Rekenen-wiskunde gr 8	6423	1473	1868	674
Spelling gr 6	7193	703	2130	412
Spelling gr 7	7707	189	2486	56
Spelling gr 8	6765	1131	2053	489

### Bepalen van de positie in het 4<sup>e</sup> jaar VO

De positie in het vierde jaar van het VO is bepaald aan de hand van een aangepaste variant op de leerjarenladder van Bosker, Van der Velden en Hofman (1985). De leerjarenladder drukt de afstand in leerjaren uit tot de top van het onderwijssysteem. Voor dit onderzoek is het maximale niveau 9: leerlingen die in vier jaar zijn gekomen tot 4-VWO. De leerjarenladder houdt rekening met zittenblijven en doorstroom. Een leerling die blijft zitten, blijft steken op de leerjarenladder. Dit zelfde geldt voor afstroom naar een lager onderwijstype. Voor gecombineerde onderwijstypen geldt een tussenliggende waarde.

	Leerjaar 1	Leerjaar 2	Leerjaar 3	Leerjaar 4
Praktijkonderwijs <sup>25</sup>	trede 1	2	3	4
VMBO-BL	2	3	4	5
VMBO-B/K	2,5	3,5		
VMBO-KL	3	4	5	6
VMBO-TL/GL	4	5	6	7
TL/HAVO	4,5	5,5		
HAVO	5	6	7	8
HAVO/VWO	5,5	6,5	7,5	
VWO	6	7	8	9

<sup>25</sup> Het praktijkonderwijs werkt niet met leerjaren. In plaats daarvan zijn verblijfsjaren als leerjaren beschouwd.

## Bijlage 2 Voorspellende waarde door de jaren heen

De vraag in dit onderzoek hoe goed scholen met de door de jaren heen verzamelde toetsscores van leerlingen kunnen voorspellen hoe zij in het VO doorstromen. Hierbij gaan we er van uit dat de school dit doet als de leerling in groep 8 zit.

Een andere invalshoek is ook denkbaar: hoe goed kun je in groep 6 voorspellen hoe een leerling in het VO doorstroomt, en hoeveel beter (of slechter) wordt die voorspelling in groep 7 of groep 8? In deze bijlage gaan we hier op in.

De toetsen in groep 6 verklaren zo'n 46,5% van de positie in het vierde VO-jaar (38% van de leerlingen komt uit op het voorspelde niveau). Met de toetsen in groep 7 daaraan toegevoegd, wordt zo'n 54% van de positie op de leerjarenladder verklaard (42% overeenkomst). Zoals we in hoofdstuk 4 al zagen, zorgen alle LOVS-toetsen tezamen voor een verklaarde variantie van .575 (44% overeenkomst).

Model	R <sup>2</sup> Oude toetsen	Percentage overeenkomst
1. BL, RW en SP groep 6	.465	38%
2. Bovenstaande + BL, RW en SP groep 7	.540*	42%
3. Bovenstaande + BL, RW en SP groep 8	.575*	44%
N= 2417		

\* de toegevoegde toetsen zorgen voor een significante toename van de proportie verklaarde variantie t.o.v. het voorgaande model

## Bijlage 3 Ongestandaardiseerde betacoëfficiënten

In onderstaande tabellen geven wij de ongestandaardiseerde betacoëfficiënten weer voor het berekenen van een “voorspeld” schooladvies of positie op de leerjarenladder in het vierde VO-jaar. Deze betacoëfficiënten gaan uit van **gestandaardiseerde** LOVS-scores (zie bijlage 1), omdat hierin wordt gecorrigeerd voor de verschillende schalen van de verschillende toetsen.<sup>26</sup>

### Schooladvies

Ongestandaardiseerde betacoëfficiënten voor het berekenen van het advies waarbij 1= VMBO-BL, 2= VMBO-KL, 3= VMBO-GL/TL, 4= HAVO, 5=VWO.

Toets	Model 1 - tabel 3.1 Oude toetsen	Model 4 - tabel 3.1 Oude toetsen	Model 1 - tabel 3.1 Nieuwe toetsen	Model 4 - tabel 3.1 Nieuwe toetsen
Constante	1.053	.571	.611	.339
Begrijpend Lezen gr .8	.028	.009	.020	.010
Rekenen-Wiskunde gr. 8	.020	.017	.032	.020
Spelling gr. 8		.003		.004
Begrijpend Lezen gr .7		.006		.010
Rekenen-Wiskunde gr. 7		.006		.004
Spelling gr. 7		.004		.004
Begrijpend Lezen gr .6		.007		.004
Rekenen-Wiskunde gr. 6		.003		.003
Spelling gr. 6		.002		.002
<b>R<sup>2</sup></b>	<b>.706</b>	<b>.761</b>	<b>.760</b>	<b>.817</b>
<b>N</b>	<b>5532</b>	<b>5532</b>	<b>1812</b>	<b>1812</b>

### Leerjarenladder

Ongestandaardiseerde betacoëfficiënten voor het berekenen van de positie op de leerjarenladder waarbij 1= PrO-1<sup>e</sup>, 9= VWO-4<sup>e</sup> etc. (zie bijlage 1).

Toets	Model 1 - tabel 4.2	Model 4 - tabel 4.2
Constante	5.033	4.553
Begrijpend Lezen gr .8	.023	.011
Rekenen-Wiskunde gr. 8	.022	.014
Spelling gr. 8		.001
Begrijpend Lezen gr .7		.009
Rekenen-Wiskunde gr. 7		.006
Spelling gr. 7		.005
Rekenen-Wiskunde gr. 6		.007
Spelling gr. 6		.002
<b>R<sup>2</sup></b>	<b>.527</b>	<b>.575</b>
<b>N</b>	<b>2417</b>	<b>2417</b>

<sup>26</sup> Uiteraard is het desgewenst ook mogelijk om dergelijke berekeningen te maken voor de ruwe (niet gestandaardiseerde) vaardigheidsscores.